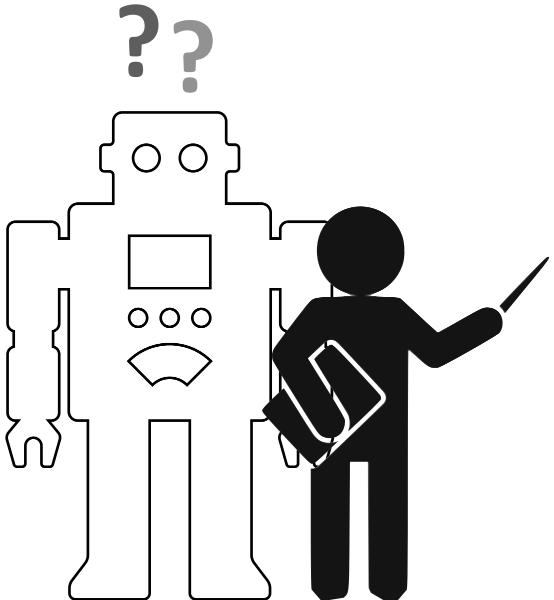
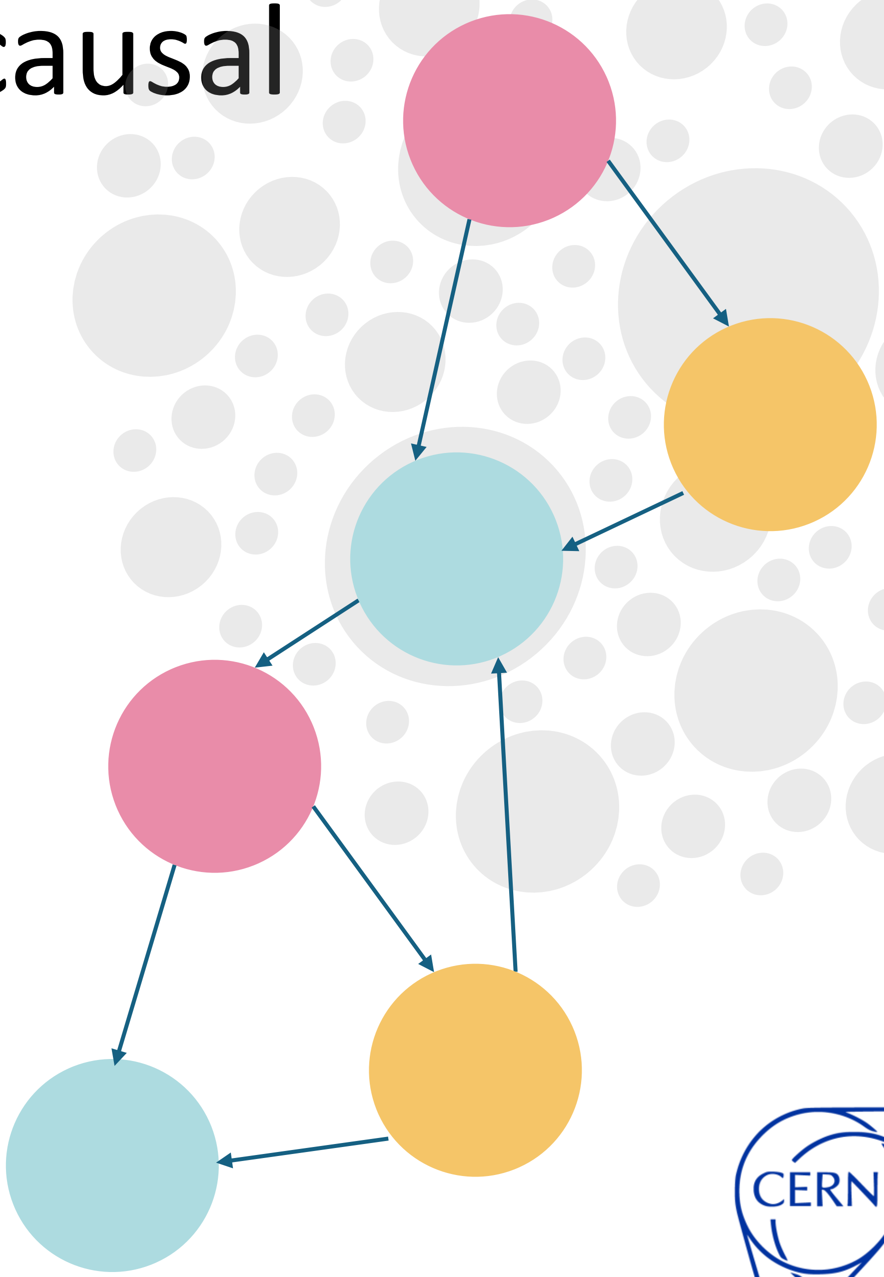
Reinforcement learning for causal discovery and data quality monitoring

Funded by Hoffman la Roche

CERN OpenLab technical Workshop

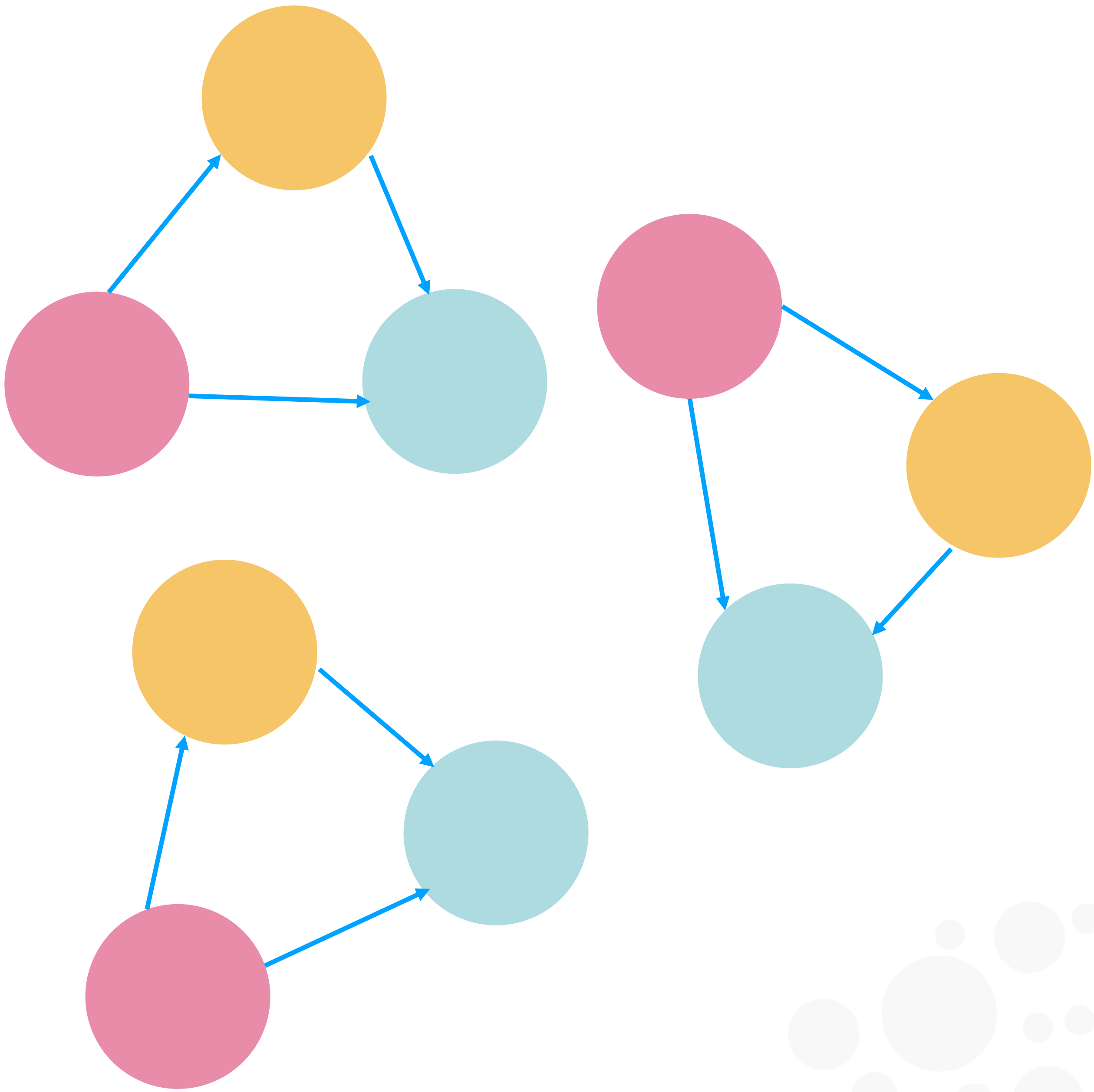
Olivia Jullian Parra (CERN, Geneva)
Sara Zoccheddu (University of Zurich, Zurich)
Shah Rukh Qasim (University of Zurich, Zurich)
Julián García Pardiñas (CERN, Geneva)
Nicola Serra (University of Zurich/CERN, Geneva)
Milo Alan Puhan (University of Zurich, Zurich)
Henock Yebyo (University of Zurich, Zurich)
Maximilian Janisch, (University of Zurich, Zurich)
Thomas Lehéricy,(University of Zurich, Zurich)





Index

Causal inference challenge



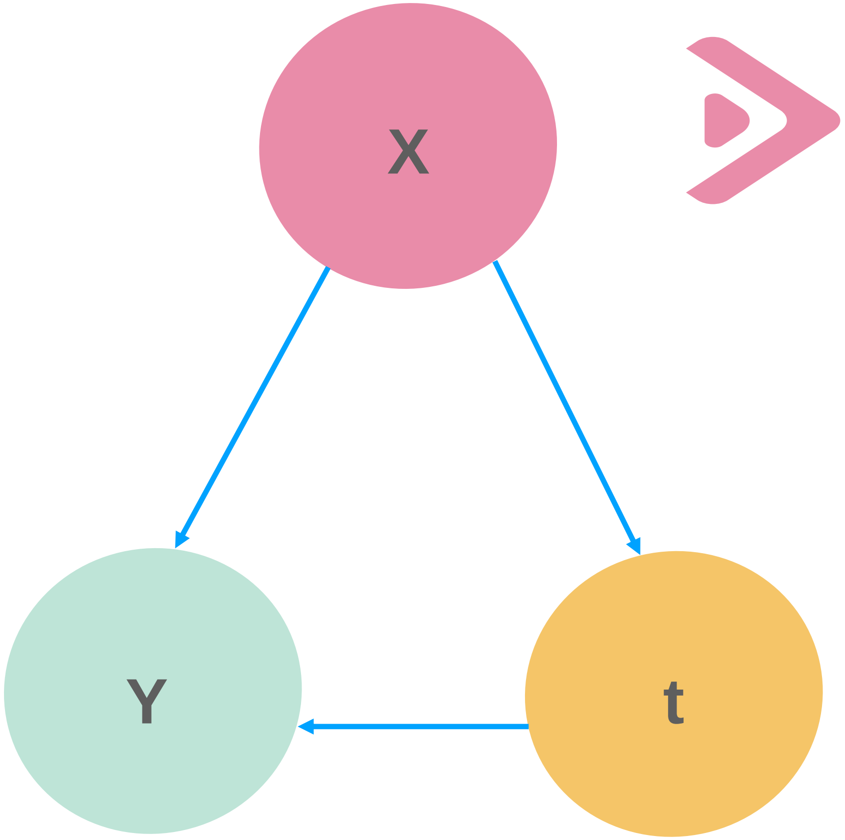
When studying a treatment for a disease…

Treatment taken

Sufficient to accept the treatment effect ?

Treatment non taken

Causal inference challenge



Confounders

Patient variables that affect the outcome and treatment decision

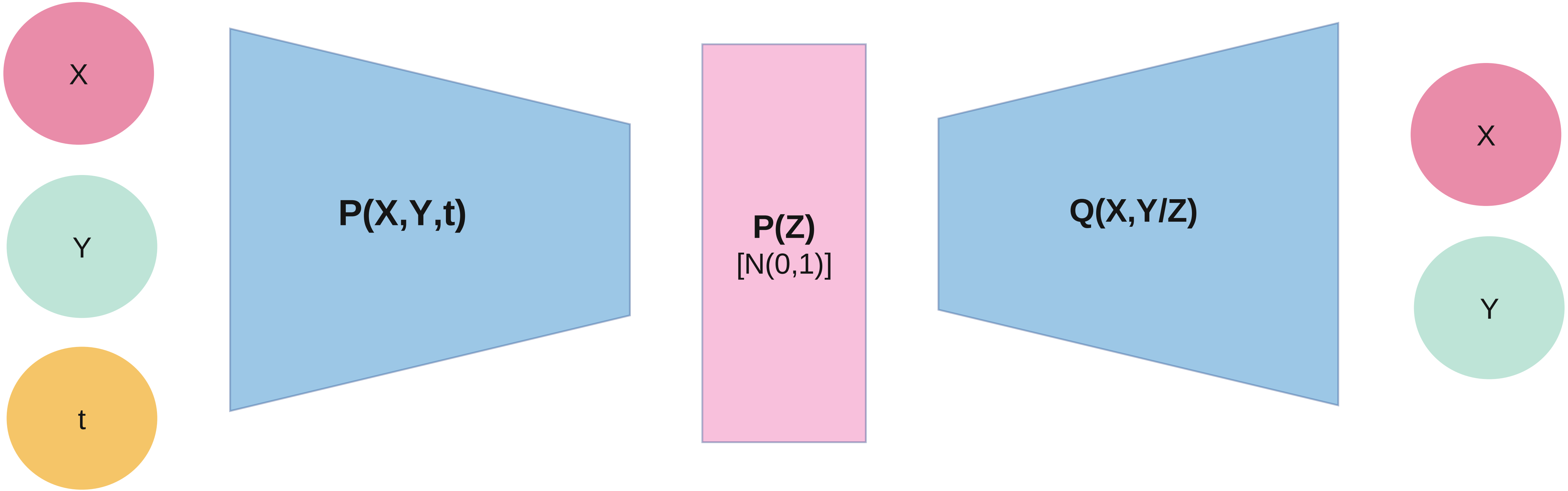
Outcome

Treatment effect on patient variables

Treatment

Treatment whose effect is under study

Machine Learning for causal inference



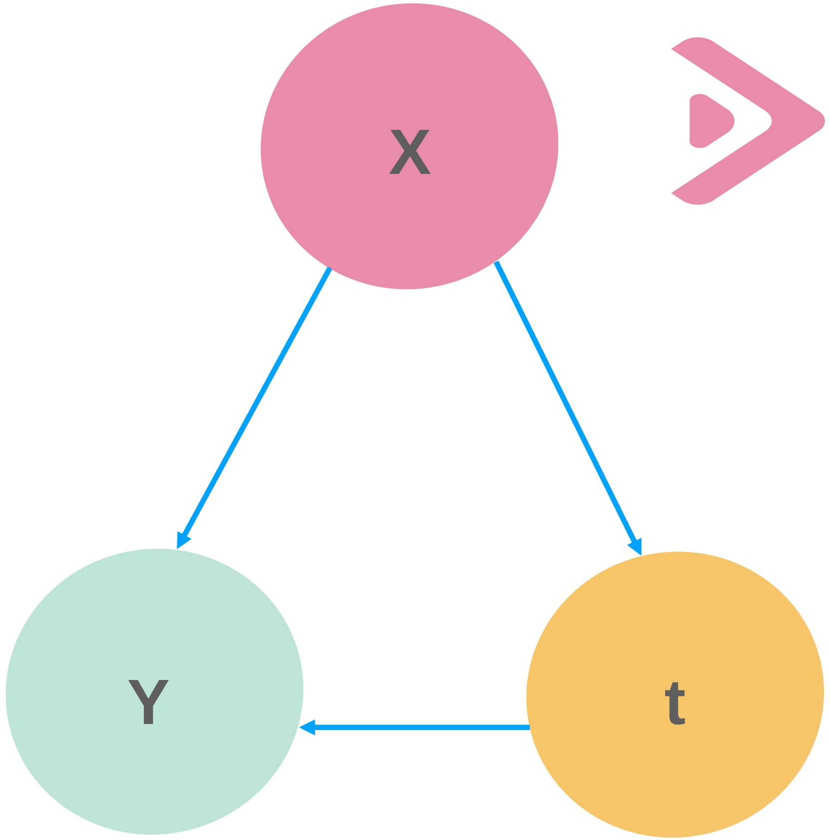
P(Z)=Q(Z/X,Y,t)

What CVAE offers

- Extract Z from X,Y,t (p(Z/X,Y,t)): acquire all the data responsible for the outcome Y
- Predict Y (q(Y/Z,X,t)): have counterfactual data to understand how t affects the outcome Y

https://indico.cern.ch/event/1225408/contributions/5249322/

Proof of concept: statins treatment effect



Confounders

30 patient variables, some of them categorical

Outcome

Cardiovascular diseases

Treatment

Statins to lower the cholesterol

Results on treatment effect prediction

CVAE

0.79

RCT

0.74

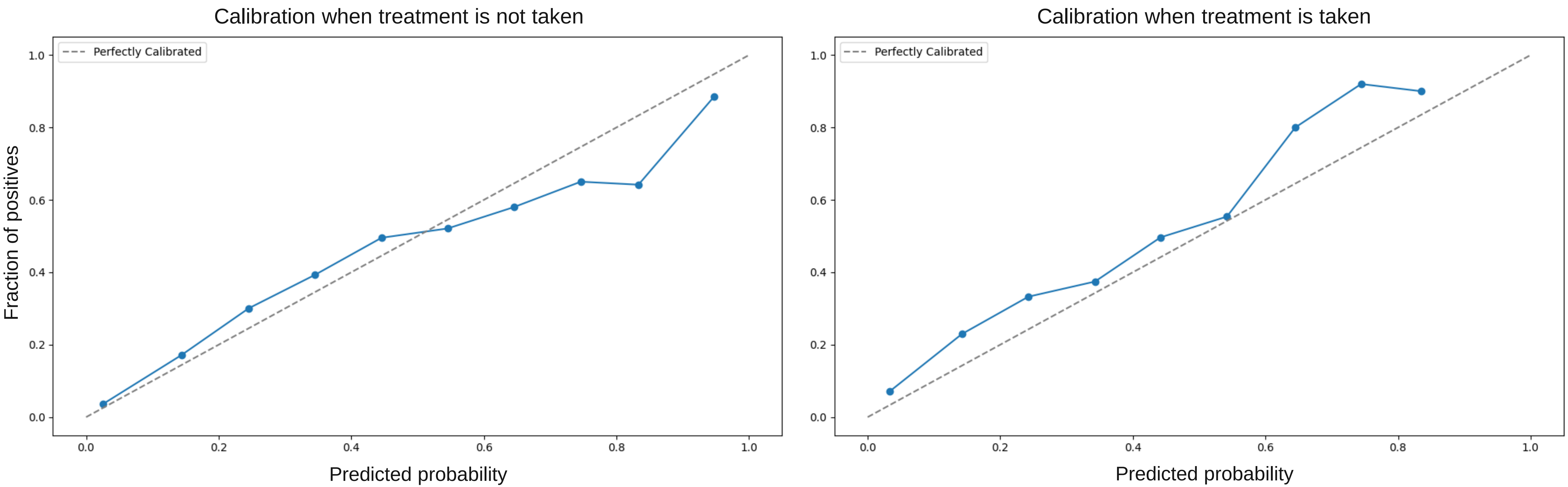
Target trial

0.78

Yebyo HG, Aschmann HE, Kaufmann M, Puhan MA. Comparative effectiveness and safety of statins as a class and of specific statins for primary prevention of cardiovascular disease: A systematic review, meta-analysis, and network meta-analysis of randomized trials with 94,283 participants. Am Heart J. 2019 Apr;210:18-28. doi: 10.1016/j.ahj.2018.12.007. Epub 2019 Jan 10. PMID: 30716508.

Yebyo, H. G., Günthard, H. F., Rehfuess, E. A., Serra, N., Haile, S. R., Senn, O., ... & Puhan, M. A. Statins for Primary Prevention of Cardiovascular Events in People Living with HIV: A Target Trial and Benefit Harm Balance Modelling Study. Available at SSRN 4427462.

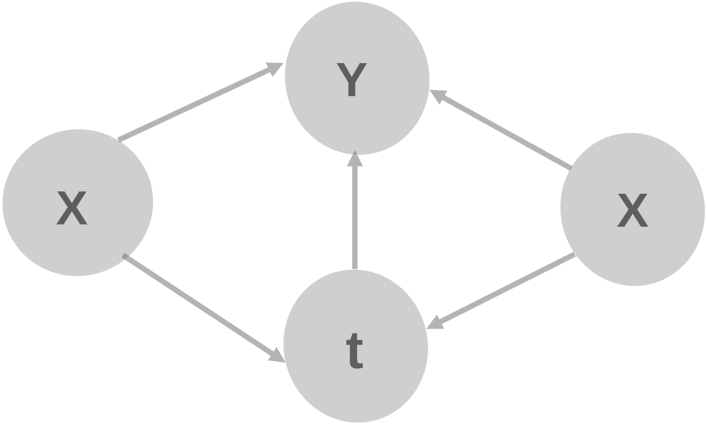
CVAE calibration



Small overestimation of the risk when the treatment is taken

Limitations

+ Unobserved confounders: what is the real X?

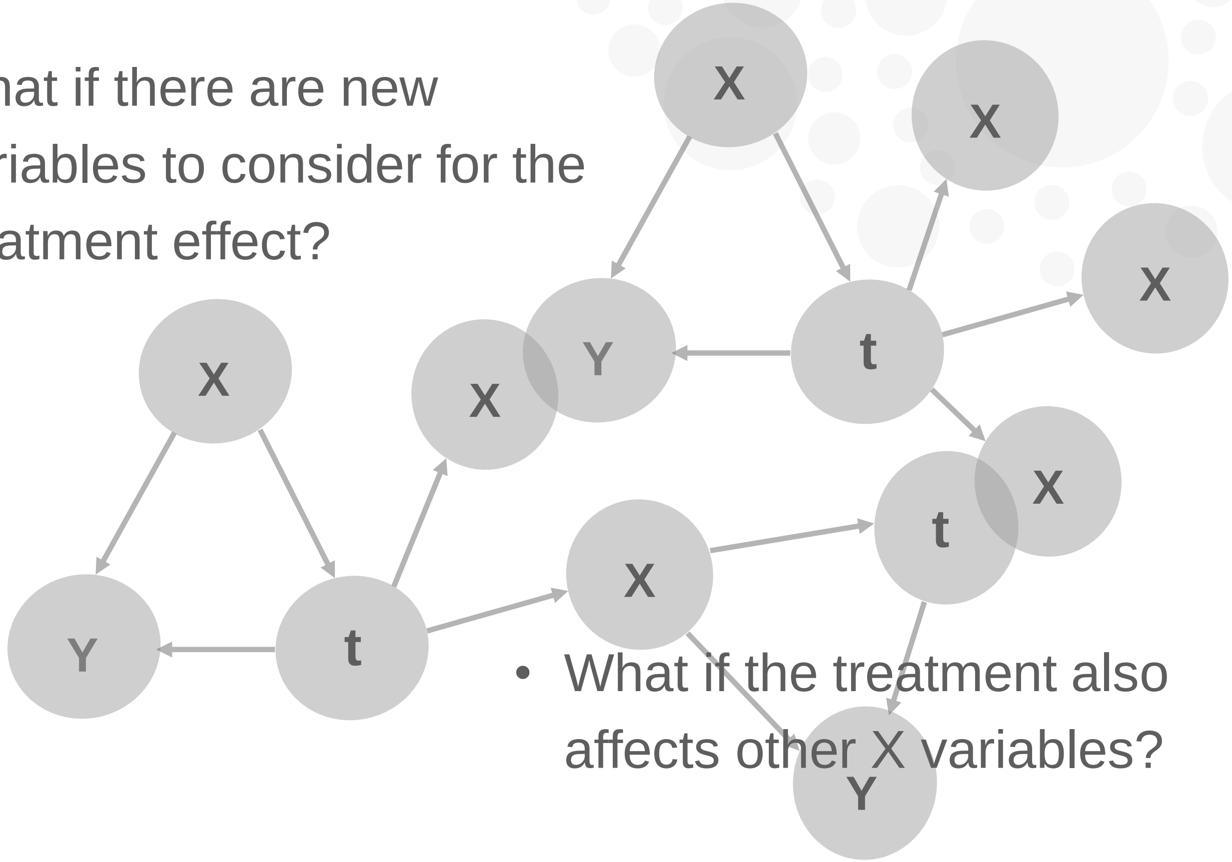


+ Multiple interactions

- What if there are multiple treatments in place to test?

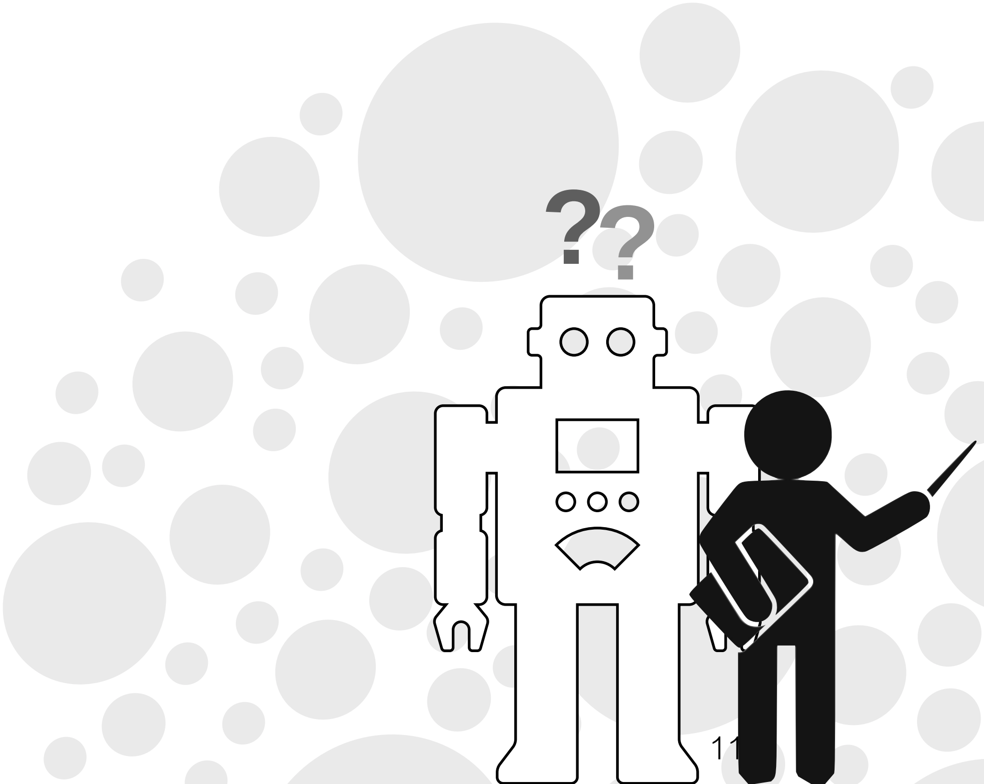
+ Benefit-Harm treatment study

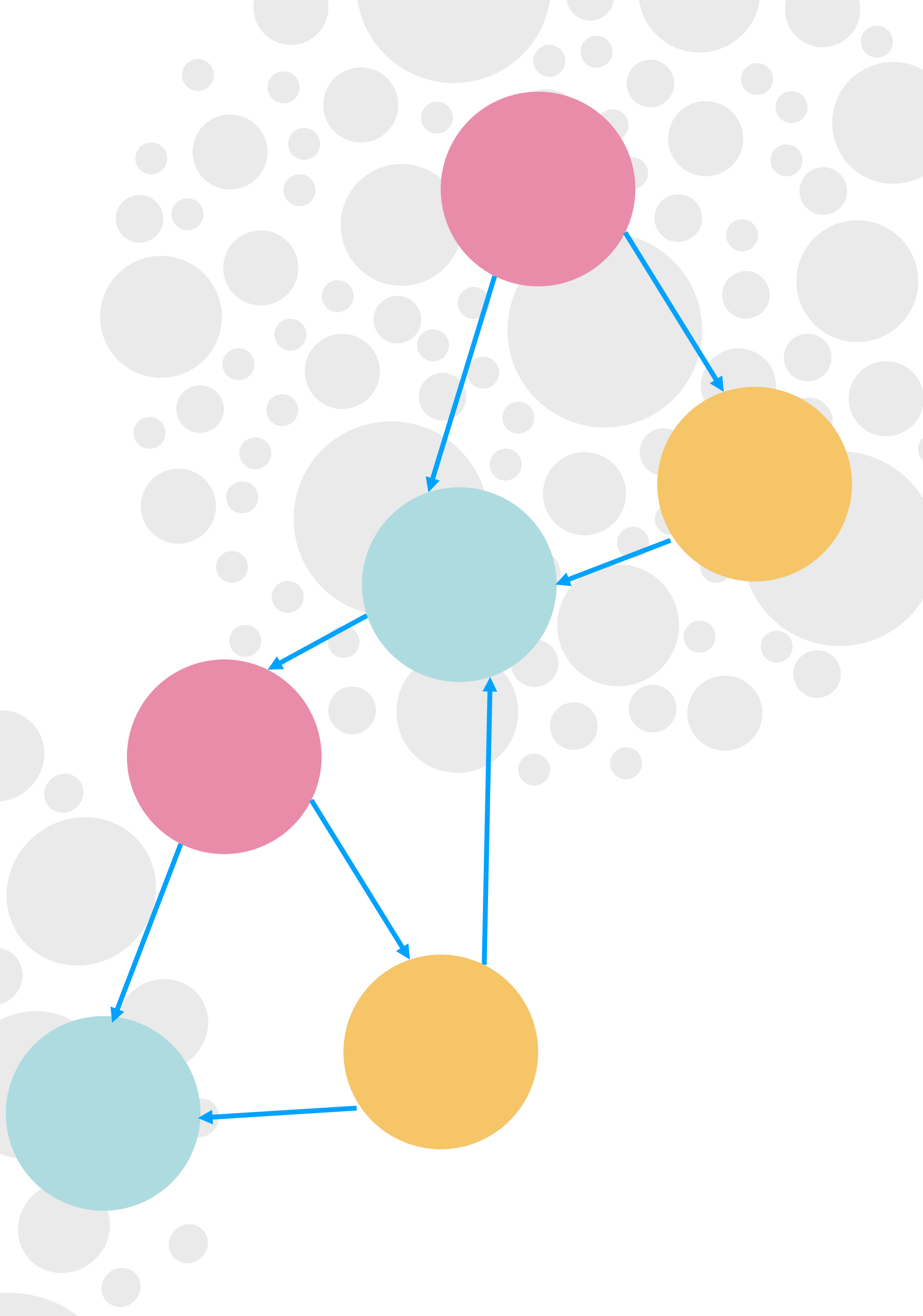
- What if there are new variables to consider for the treatment effect?



- What if the treatment also affects other X variables?

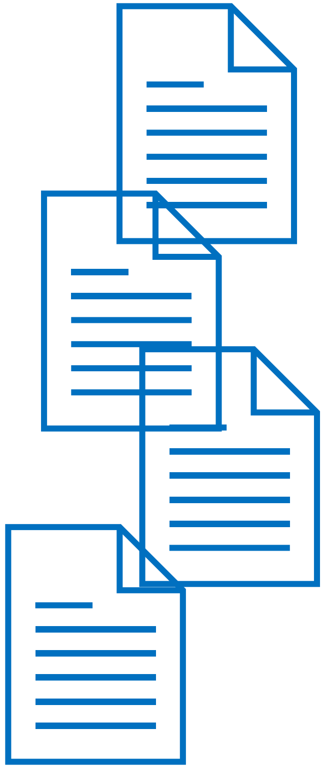
Causal discovery

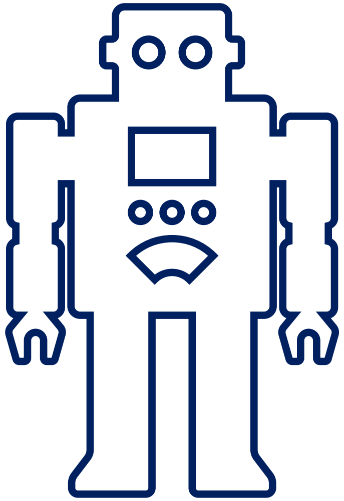




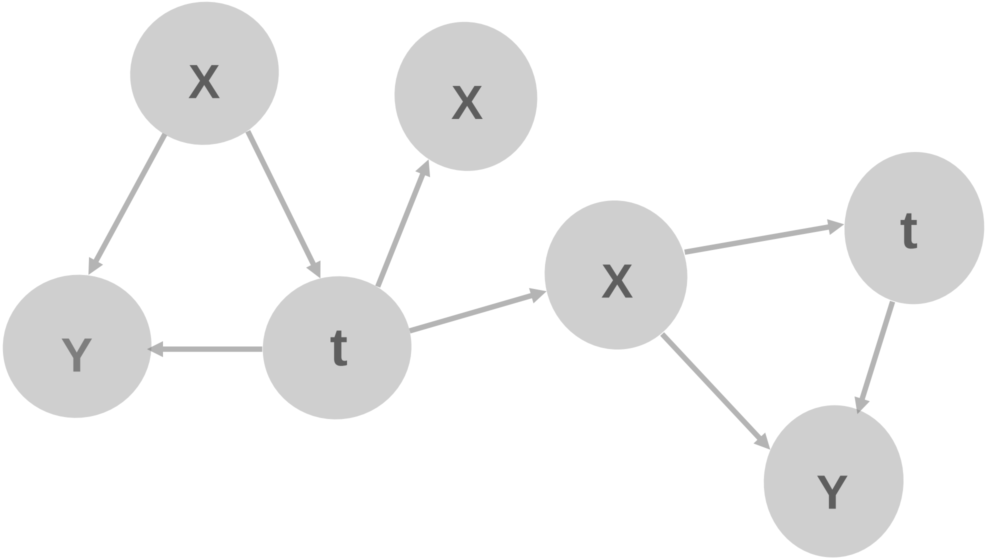
Causal Discovery pipeline

We have observational data



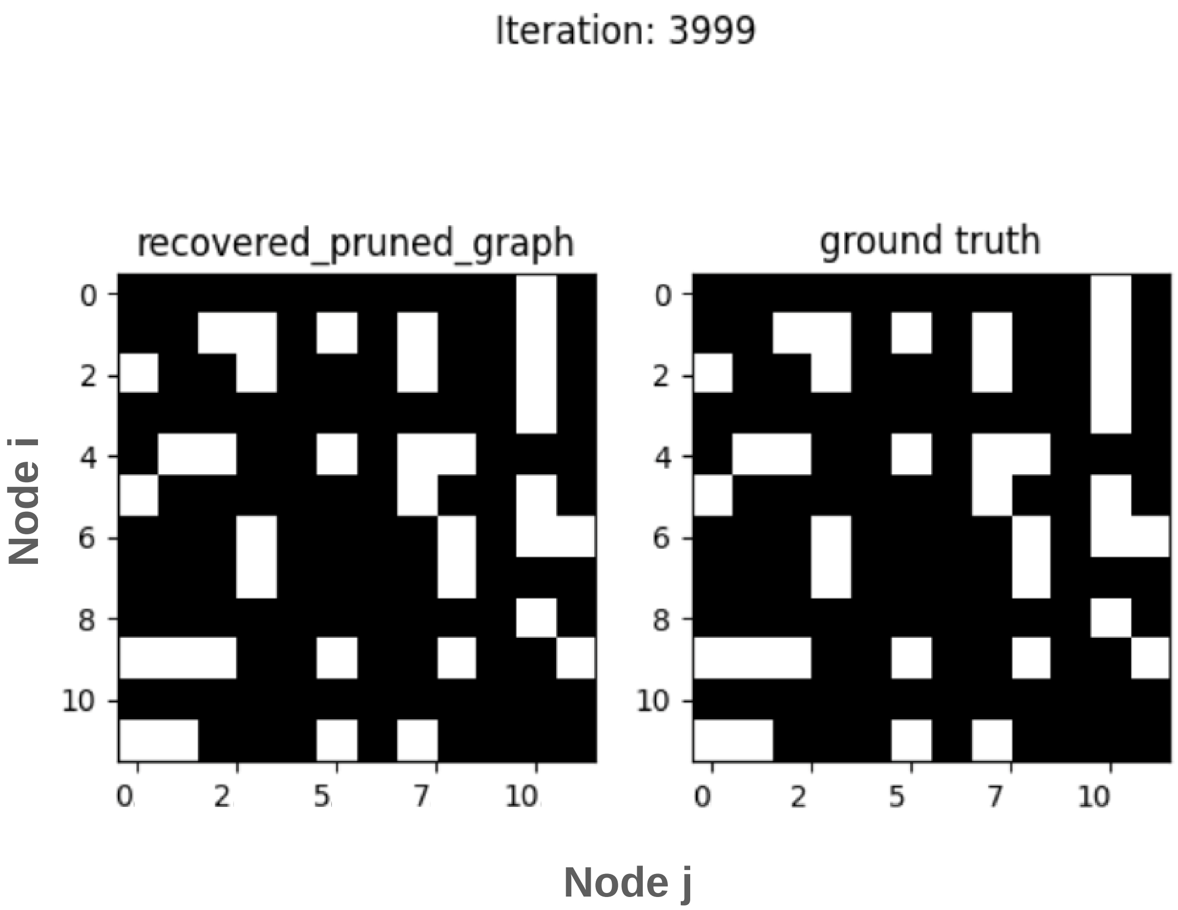


We find causal relationships



We use a Reinforcement Learning (RL) algorithm

Experiments



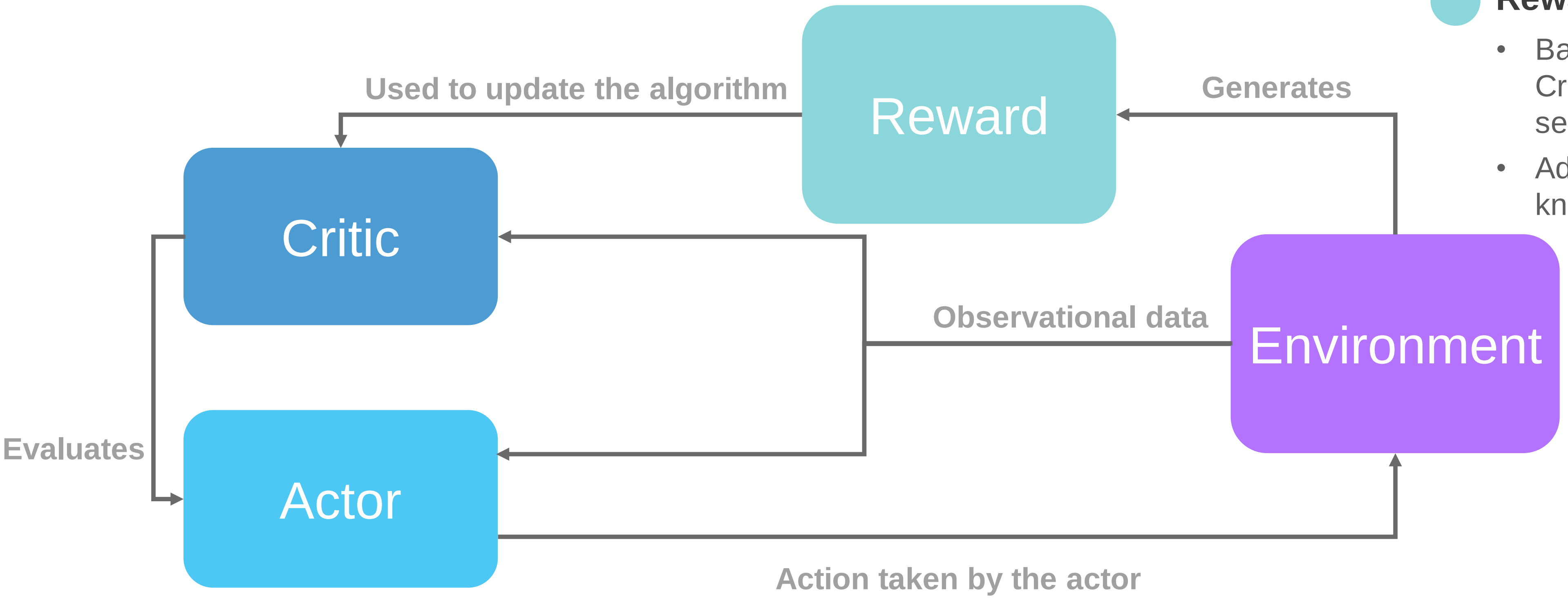
Toy dataset

- Variables have linear-gaussian dependencies

- The variance of the values is the same for all of them

Zhu, S., Ng, I., & Chen, Z. (2019). Causal discovery with reinforcement learning. arXiv preprint arXiv:1906.04477.

Actor-Critic model for RL



Reward

- Based on the Bayesian Information Criterion for the best graph selection
- Adds penalties to include expert knowledge

Actor

- Takes a state derived from all the observational data
- Works under a policy
- Proposes a causal graph candidate

Critic

- Evaluates the graphs proposed by the actor
- Predicts again the obs. data using the graph proposed (needs optimization)

Environment

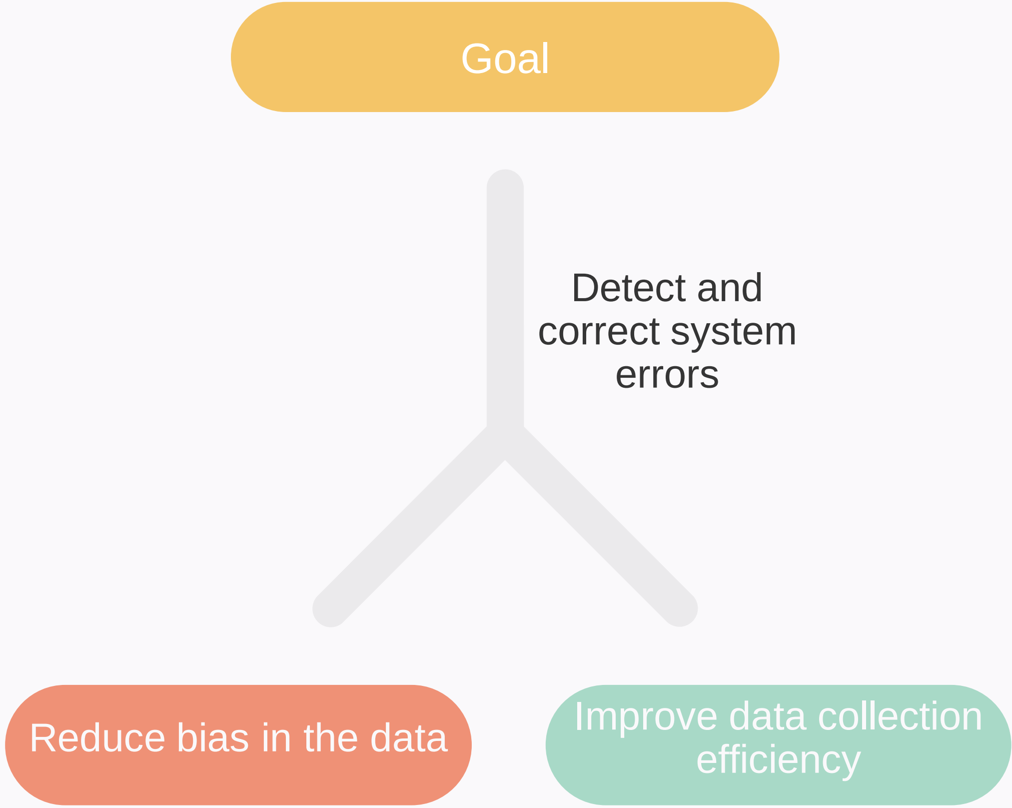
- Representation of the observational data collection
- Represents all the possible patient variables, treatments and outcomes

Data quality monitoring

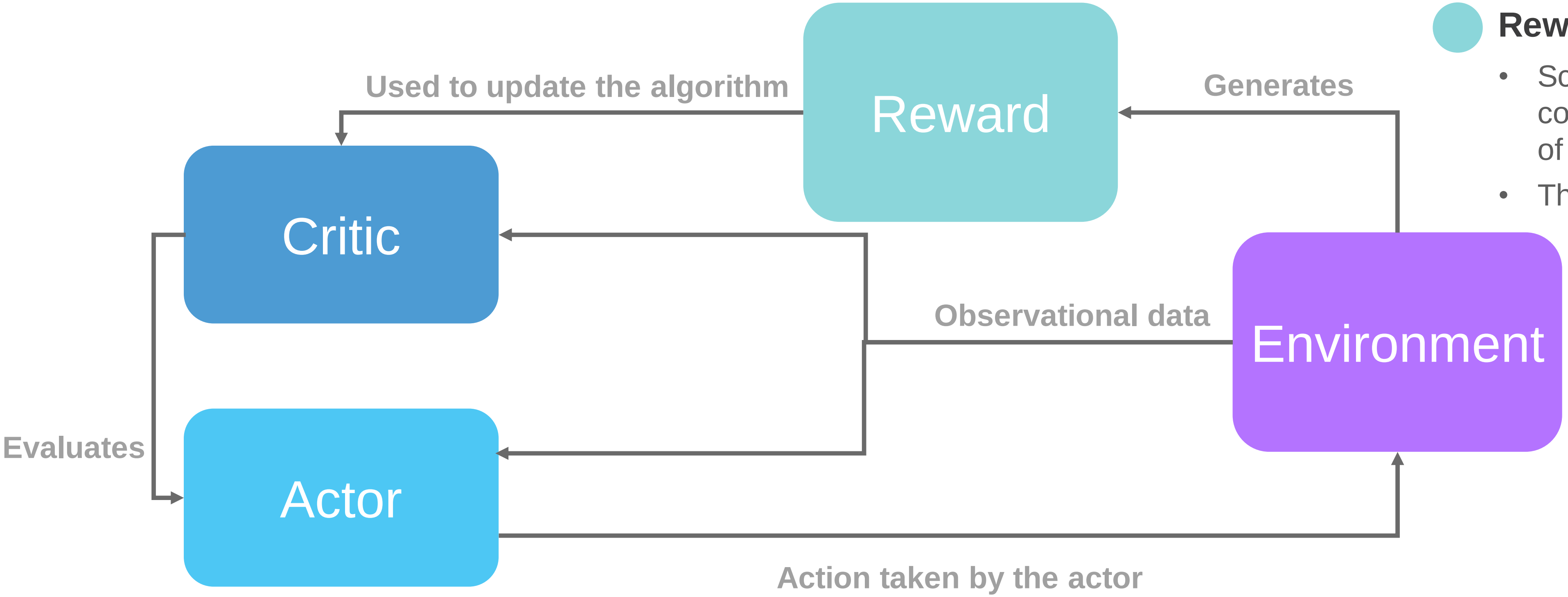
Systems are imperfect and may bias the collected data.

When the system status is incorrectly classified as good, measurements are biased.

When the system status is incorrectly classified as bad, data collection efficiency reduces.



Actor-Critic model with human feedback



Reward

- Scheme reward based to allow the correctness and confidence level of the actor
- The human will provide the reward

Actor

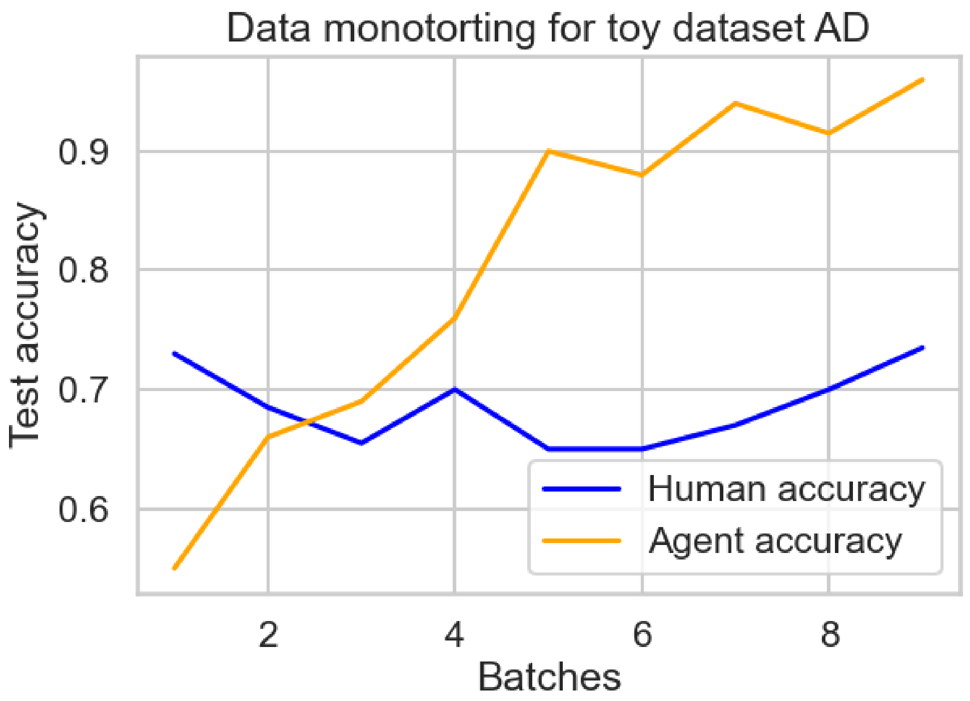
- Proposes an action for the system status and tries to integrate the human feedback to its policy

Environment

- Representation of the system monitoring
- The state is a collected histogram of the system status

https://indico.cern.ch/event/1297159/contributions/5729221/

Results



Probability of human failure: 0.3

- We can **achieve superhuman condition** for the offline regime
- **How do we know when to rely on the machine?** We are currently performing studies on the detection of superhuman condition during training, without having access to the ground truth[1]

[1] Bar, O., Drory, A., & Giryes, R. (2022, May). A spectral perspective of DNN robustness to label noise. In International Conference on Artificial Intelligence and Statistics (pp. 3732-3752). PMLR.

Conclusions

Causal inference achieve. We can use them for more complex use cases to complement RCTs

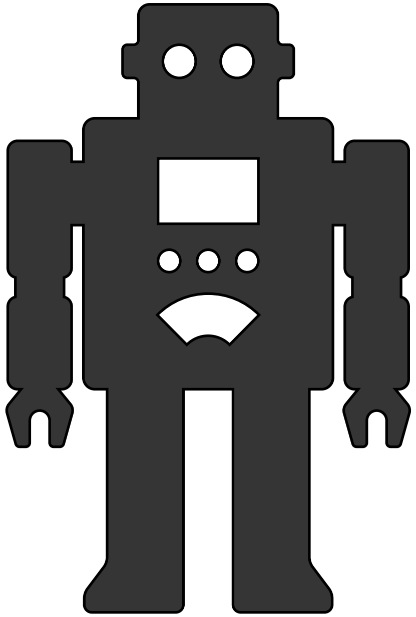
Causal Inference

First steps done towards causal discovery. Able to abstract the algorithm and adapted it to new use cases

Causal discovery

Ability to assist the shifter decisions without necessity of their constant feedback and automatize some system’s checks

Data monitoring



Thanks
for your attention

Q&A

olivia.jullian.parra@cern.ch